Extremos condicionados: el método de los multiplicadores de Lagrange

(Un esquema de tres vías de utilización de la hessiana [existen más] – Muy brevemente)

(Precedido de otro uso de la hessiana para extremos no condicionados y de conceptos necesarios para su entendimiento)

Derivadas parciales sucesivas

Sean $A \subseteq \mathbb{R}^n$, A abierto, $\vec{a} \in A$, $\vec{f} : A \longrightarrow \mathbb{R}^m$ un campo vectorial y $f : A \longrightarrow \mathbb{R}$ un campo escalar.

Derivadas parciales segundas de un campo escalar f.— Las derivadas parciales de un campo escalar f pueden ser a su vez derivables; se conocen como derivadas parciales segundas o de orden dos de f.

Por ejemplo, si $n = 2$ y utilizando la notación de Leibniz, son:

$$\frac{\partial^2 f}{\partial x^2}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}(x, y)\right), \qquad \frac{\partial^2 f}{\partial x\, \partial y}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x, y)\right),$$
$$\frac{\partial^2 f}{\partial y\, \partial x}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x, y)\right), \qquad \frac{\partial^2 f}{\partial y^2}(x, y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}(x, y)\right),$$

siendo también frecuentes en la literatura otras notaciones; por ejemplo, $\dfrac{\partial^2 f}{\partial y\, \partial x}$ también se nota $\partial_{yx} f$, f_{xy} y $(f_x)_y$.

La situación puede repetirse para órdenes mayores y podemos hablar de derivadas de orden tres, cuatro, etc.

Campo escalar de clase C^n en A.— El campo escalar f es:

- de clase C^1 en A precisamente si es derivable en A y sus derivadas parciales de primer orden son continuas en A;
- de clase C^n en A ($n \geq 2$) precisamente si es n veces derivable en A, es de clase C^{n-1} en A y sus derivadas parciales de orden n son continuas en A.

Campo vectorial de clase C^n en A.— El campo vectorial $\vec{f}$ es de clase C^n en A ($n \geq 1$) precisamente si lo son en A todas sus funciones componentes.

Campo de clase C^∞ en A.— Aquél que es continuo en A, existen todas sus derivadas parciales de cualquier orden y todas éstas son funciones continuas en A.

Teorema de Schwarz para las derivadas cruzadas.— Si f es tal que existen $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial f}{\partial x_j}$ y $\dfrac{\partial^2 f}{\partial x_i\, \partial x_j}$ en $\vec{a}$ y además $\dfrac{\partial^2 f}{\partial x_i\, \partial x_j}$ es continua en $\vec{a}$, entonces también existe $\dfrac{\partial^2 f}{\partial x_j\, \partial x_i}$ en $\vec{a}$ y se satisface

$$\frac{\partial^2 f}{\partial x_j\, \partial x_i}(\vec{a}) = \frac{\partial^2 f}{\partial x_i\, \partial x_j}(\vec{a}).$$

Observación.— Un campo de clase C^2 satisface el teorema de Schwarz.

Matriz hessiana de un campo escalar $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ con derivadas parciales de segundo orden en $\vec{a}$.— Es la matriz $n \times n$ definida por

$$Hf(\vec{a}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\vec{a}) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{a}) & \dfrac{\partial^2 f}{\partial x_2^2}(\vec{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(\vec{a}) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(\vec{a}) \end{pmatrix}.$$

Por ejemplo, si $n = 2$ y si $n = 3$, con la notación habitual, $\vec{x} = (x, y)$ y $\vec{x} = (x, y, z)$, respectivamente, son

$$Hf(a_1, a_2) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a_1, a_2) & \dfrac{\partial^2 f}{\partial x \partial y}(a_1, a_2) \\ \dfrac{\partial^2 f}{\partial y \partial x}(a_1, a_2) & \dfrac{\partial^2 f}{\partial y^2}(a_1, a_2) \end{pmatrix},$$

$$Hf(a_1, a_2, a_3) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial x \partial y}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial x \partial z}(a_1, a_2, a_3) \\ \dfrac{\partial^2 f}{\partial y \partial x}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial y^2}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial y \partial z}(a_1, a_2, a_3) \\ \dfrac{\partial^2 f}{\partial z \partial x}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial z \partial y}(a_1, a_2, a_3) & \dfrac{\partial^2 f}{\partial z^2}(a_1, a_2, a_3) \end{pmatrix}.$$

Hessiano.— Es el determinante de la matriz hessiana.

Polinomio homogéneo.— Polinomio cuyos términos no nulos tienen todos el mismo grado. Por ejemplo, $x^6 + x^3y^3 + x^4y^2 - xy^5$, pues todos tienen grado 6 ($6 = 3 + 3 = 4 + 2 = 1 + 5$).

Forma cuadrática.— Polinomio homogéneo de grado 2.

Forma cuadrática asociada a una matriz.— Toda matriz cuadrada $A = \left(a_{ij}\right)_{n \times n}$ determina una forma cuadrática $q_A : \mathbb{R}^n \longrightarrow \mathbb{R}$ definida para $\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ por

$$q_A(\vec{x}) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

esto es, por

$$q_A(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j.$$

Se dice que la forma cuadrática es: **definida positiva** precisamente si todos los menores principales[1] de A son positivos; **definida negativa** precisamente si son negativos todos los menores principales de A de índice impar y son positivos todos los de índice par; **semidefinida positiva** precisamente si todos los menores principales de A son no negativos (≥ 0); **semidefinida negativa** precisamente si son no positivos (≤ 0) todos los menores principales de A de índice impar y son no negativos (≥ 0) todos los de índice par, e **indefinida** en otros casos.

Función hessiana de un campo escalar $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ con derivadas parciales de segundo orden en $\vec{a}$.— Es la forma cuadrática $q_{Hf(\vec{a})} : \mathbb{R}^n \longrightarrow \mathbb{R}$ determinada por $Hf(\vec{a})$, esto es, aplicada a $\vec{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$ está definida por

$$q_{Hf(\vec{a})}(\vec{h}) = \begin{pmatrix} h_1 & \cdots & h_n \end{pmatrix} Hf(\vec{a}) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix},$$

es decir, por

$$q_{Hf(a_1, \ldots, a_n)}(h_1, \ldots, h_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \, \partial x_j}(a_1, \ldots, a_n) h_i h_j.$$

Por ejemplo, si $n = 2$ y si $n = 3$, con la notación habitual, $\vec{x} = (x, y)$ y $\vec{x} = (x, y, z)$, respectivamente, son

$$qf_{Hf(a_1, a_2)}(h_1, h_2) = \frac{\partial^2 f}{\partial x^2}(a_1, a_2) h_1^2 + \frac{\partial^2 f}{\partial y^2}(a_1, a_2) h_2^2 + 2 \frac{\partial^2 f}{\partial x \, \partial y}(a_1, a_2) h_1 h_2,$$

[1] Dada una matriz A, un menor, esto es, un determinante de una submatriz de A, se dice que es un **menor principal** precisamente si el conjunto de los índices de sus filas es el mismo que el conjunto de los índices de sus columnas; por ejemplo, si $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, sus menores principales son: $\left|a_{11}\right|$, $\left|a_{22}\right|$, $\left|a_{33}\right|$, $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, $\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$, $\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ y $|A|$.

$$q f_{Hf(a_1,a_2,a_3)}(h_1,h_2,h_3) = \frac{\partial^2 f}{\partial x^2}(a_1,a_2,a_3)h_1^2 + \frac{\partial^2 f}{\partial y^2}(a_1,a_2,a_3)h_2^2 + \frac{\partial^2 f}{\partial z^2}(a_1,a_2,a_3)h_3^2 + 2\frac{\partial^2 f}{\partial x\,\partial y}(a_1,a_2,a_3)h_1h_2 + 2\frac{\partial^2 f}{\partial x\,\partial z}(a_1,a_2,a_3)h_1h_3 + 2\frac{\partial^2 f}{\partial y\,\partial z}(a_1,a_2,a_3)h_2h_3.$$

Diferenciabilidad de un campo vectorial

Sean $A \subseteq \mathbb{R}^n$, A abierto, $\vec{a} \in A$, $\vec{f} : A \longrightarrow \mathbb{R}^m$ un campo vectorial, $f : A \longrightarrow \mathbb{R}$ un campo escalar y $\vec{u} \in \mathbb{R}^n$, unitario.

Campo vectorial diferenciable en un punto $\vec{a}$.— El campo vectorial $\vec{f}$ es diferenciable en $\vec{a}$ precisamente si existe una aplicación lineal[2] $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ tal que

$$\lim_{\vec{x}\to\vec{a}} \frac{\vec{f}(\vec{x}) - \vec{f}(\vec{a}) - L(\vec{x}-\vec{a})}{\|\vec{x}-\vec{a}\|} = \vec{0}.$$

Diferencial de $\vec{f}$ en $\vec{a}$.— Si $\vec{f}$ es diferenciable en $\vec{a}$, la aplicación lineal L es única; esta aplicación lineal se conoce como diferencial de $\vec{f}$ en $\vec{a}$ y se representa como $d\vec{f}(\vec{a})$.

Campo escalar diferenciable en un punto.— El campo escalar f es diferenciable en $\vec{a}$ precisamente si existe una aplicación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que

$$\lim_{\vec{x}\to\vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - L(\vec{x}-\vec{a})}{\|\vec{x}-\vec{a}\|} = \vec{0}.$$

Diferencial de un campo escalar en un punto.— Si f es diferenciable en $\vec{a}$, la aplicación lineal L es única; esta aplicación lineal se conoce como diferencial de f en $\vec{a}$ y se representa como $df(\vec{a})$.

Campo vectorial diferenciable en un conjunto abierto A.— El campo vectorial $\vec{f}$ es diferenciable en A precisamente si lo es en todo punto $\vec{a}$ de A.

Una condición necesaria de diferenciabilidad.— Si $\vec{f}$ es diferenciable en $\vec{a}$, entonces $\vec{f}$ es continua en $\vec{a}$.

Otra condición necesaria de diferenciabilidad.— Si $\vec{f}$ es diferenciable en $\vec{a}$, entonces existen todas las derivadas direccionales de $\vec{f}$ en $\vec{a}$, $D_{\vec{u}}\vec{f}(\vec{a})$, con $\vec{u}$ un vector unitario; además, se satisface que $D_{\vec{u}}\vec{f}(\vec{a}) = d\vec{f}(\vec{a})(\vec{u})$.

Otra condición necesaria de diferenciabilidad.— Si $\vec{f}$ es diferenciable en $\vec{a}$, entonces existen todas las derivadas parciales de primer orden de $\vec{f}$ en $\vec{a}$, ya que $\dfrac{\partial \vec{f}}{\partial x_i}(\vec{a}) = D_{\vec{e}_i}\vec{f}(\vec{a})$, con $\vec{e}_i$ el vector unitario $(0,\dots,0,\underbrace{1}_{i},0,\dots,0)$ —en otras palabras, $\dfrac{\partial \vec{f}}{\partial x_i}(\vec{a}) = d\vec{f}(\vec{a})(\vec{e}_i)$; y, en otras, $\dfrac{\partial \vec{f}}{\partial x_i}(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{e}_i$ (donde $\cdot$ es el producto escalar[3])—.

Una condición suficiente de diferenciabilidad.— Si $\vec{f}$ es de clase C^1 en A, entonces $\vec{f}$ es diferenciable en A.

Matriz jacobiana de un campo vectorial $\vec{f} = (f_1, f_2, \dots, f_m)$ diferenciable en $\vec{a}$.— Es la matriz $m \times n$ asociada a la diferencial de $\vec{f}$ en $\vec{a}$ como aplicación lineal $d\vec{f}(\vec{a}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, es decir,

$$J\vec{f}(\vec{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \frac{\partial f_2}{\partial x_1}(\vec{a}) & \frac{\partial f_2}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \frac{\partial f_m}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{pmatrix},$$

[2] Sean V y W espacios vectoriales, ambos sobre un mismo cuerpo K; se dice que $L : V \longrightarrow W$ es una **aplicación lineal** precisamente si $\forall \vec{u}, \vec{v} \in V$, $\forall k \in K$, se satisface $L(u+v) = L(u) + L(v)$ y $L(ku) = kL(u)$.

[3] El **producto escalar** de dos vectores reales $\vec{u}(u_1, u_2, \cdots, u_n), \vec{v}(v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$ se define por $\vec{u} \cdot \vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$.

que reescrita en términos del gradiente[4] queda

$$J\vec{f}(\vec{a}) = \begin{pmatrix} \nabla f_1(\vec{a}) \\ \nabla f_2(\vec{a}) \\ \vdots \\ \nabla f_m(\vec{a}) \end{pmatrix}.$$

Diferencial de un campo vectorial en un punto (bis).— Si $\vec{f}$ es diferenciable en $\vec{a} \in A$, entonces la diferencial de $\vec{f}$ en $\vec{a}$, esto es, la aplicación lineal $d\vec{f}(\vec{a}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, aplicada a $\vec{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$, está definida por

$$d\vec{f}(\vec{a})(\vec{h}) = J\vec{f}(\vec{a}) \cdot \vec{h}.$$

Cálculo de la derivada direccional de un campo vectorial $\vec{f}$ en $\vec{a}$ según la dirección de $\vec{u}$.— Si $\vec{f}$ es un campo vectorial, como $D_{\vec{u}}\vec{f}(\vec{a}) = d\vec{f}(\vec{a})(\vec{u})$, entonces

$$D_{\vec{u}}\vec{f}(\vec{a}) = J\vec{f}(\vec{a}) \cdot \vec{u}.$$

Diferencial de un campo escalar en un punto (bis).— Si f es diferenciable en $\vec{a} \in A$, entonces la diferencial de f en $\vec{a}$, esto es, la aplicación lineal $df(\vec{a}) : \mathbb{R}^n \longrightarrow \mathbb{R}$, aplicada a $\vec{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$, está definida por

$$df(\vec{a})(\vec{h}) = \nabla f(\vec{a}) \cdot \vec{h},$$

esto es,

$$df(\vec{a})(\vec{h}) = \frac{\partial f}{\partial x_1}(\vec{a})h_1 + \frac{\partial f}{\partial x_2}(\vec{a})h_2 + \cdots + \frac{\partial f}{\partial x_n}(\vec{a})h_n.$$

Cálculo de la derivada direccional de un campo escalar f en $\vec{a}$ según la dirección de $\vec{u}$.— Si f es un campo escalar, como $D_{\vec{u}}f(\vec{a}) = df(\vec{a})(\vec{u})$, entonces

$$D_{\vec{u}}f(\vec{a}) = \nabla f(\vec{a}) \cdot \vec{u},$$

o lo que es igual,

$$D_{\vec{u}}f(\vec{a}) = \|\nabla f(\vec{a})\| \cos\varphi,$$

con φ el ángulo que forman $\nabla f(\vec{a})$ y $\vec{u}$.

Valores mínimo y máximo de la derivada direccional de un campo escalar f en $\vec{a}$.—

- El valor máximo de la derivada direccional (o sinónimamente, la máxima derivada direccional) de f en $\vec{a}$ es $\|\nabla f(\vec{a})\|$ y se alcanza cuando $\vec{a}$ tiene la dirección y sentido de $\nabla f(\vec{a})$.
- El valor mínimo de la derivada direccional (o sinónimamente, la mínima derivada direccional) de f en $\vec{a}$ es $-\|\nabla f(\vec{a})\|$ y se alcanza cuando $\vec{a}$ tiene la dirección de $\nabla f(\vec{a})$ pero su sentido es el contrario.
- Toda derivada direccional de f en $\vec{a}$ satisface $-\|\nabla f(\vec{a})\| \leq D_{\vec{u}}f(\vec{a}) \leq \|\nabla f(\vec{a})\|$.

Campo escalar diferenciable en un punto (bis).— El campo escalar f es diferenciable en $\vec{a}$ precisamente si existe una aplicación lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - L(\vec{h})}{\|\vec{h}\|} = \vec{0}$$

con $\vec{h} = (h_1, \dots, h_n) \in \mathbb{R}^n$, y que viene dada por

$$L(\vec{h}) = df(\vec{a})(\vec{h}),$$

esto es, por

$$df(\vec{a})(\vec{h}) = \frac{\partial f}{\partial x_1}(\vec{a})h_1 + \frac{\partial f}{\partial x_2}(\vec{a})h_2 + \cdots + \frac{\partial f}{\partial x_n}(\vec{a})h_n,$$

[4] El **gradiente** de un campo escalar $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ con derivadas parciales de primer orden en $\vec{a}$, es el campo vectorial $\nabla f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ definido para $\vec{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ por $\nabla f(\vec{a}) = \left(\frac{\partial f}{\partial x_1}(\vec{a}), \frac{\partial f}{\partial x_2}(\vec{a}), \cdots, \frac{\partial f}{\partial x_n}(\vec{a})\right)$, conociéndose éste como **vector gradiente**.

por lo que el campo escalar f es diferenciable en $\vec{a}$ precisamente si

$$\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - \dfrac{\partial f}{\partial x_1}(\vec{a})h_1 - \dfrac{\partial f}{\partial x_2}(\vec{a})h_2 - \cdots - \dfrac{\partial f}{\partial x_n}(\vec{a})h_n}{\|\vec{h}\|} = \vec{0}.$$

Extremos para campos escalares

Mínimo absoluto (o global).— Siendo $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$, se dice que f alcanza/presenta/tiene un mínimo absoluto (o global) en $\vec{a}$ precisamente si para todo $\vec{x} \in A$, $f(\vec{x}) \geq f(\vec{a})$. Este mínimo absoluto se dice que es estricto si lo que sucede es que $f(\vec{x}) > f(\vec{a})$.

Máximo absoluto (o global).— Siendo $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$, se dice que f alcanza/presenta/tiene un máximo absoluto (o global) en $\vec{a}$ precisamente si para todo $\vec{x} \in A$, $f(\vec{x}) \leq f(\vec{a})$. Este máximo absoluto se dice que es estricto si lo que sucede es que $f(\vec{x}) < f(\vec{a})$.

Extremo absoluto (o global).— Un punto donde un campo alcance un mínimo o un máximo absoluto.

Mínimo local (o relativo).— Siendo $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$ un punto no aislado de A, se dice que f alcanza/presenta/tiene un mínimo local (o relativo) en $\vec{a}$ precisamente si existe una bola abierta B centrada en $\vec{a}$ tal que para todo $\vec{x} \in B \cap A$, $f(\vec{x}) \geq f(\vec{a})$. Este mínimo local se dice que es estricto si lo que sucede es que $f(\vec{x}) > f(\vec{a})$.

Máximo local (o relativo).— Siendo $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$ un punto no aislado de A, se dice que f alcanza/presenta/tiene un máximo local (o relativo) en $\vec{a}$ precisamente si existe una bola abierta B centrada en $\vec{a}$ tal que para todo $\vec{x} \in B \cap A$, $f(\vec{x}) \leq f(\vec{a})$. Este máximo local se dice que es estricto si lo que sucede es que $f(\vec{x}) < f(\vec{a})$.

Extremo local (o relativo).— Un punto donde un campo alcance un mínimo o un máximo local.

Teorema de Weierstrass.— Si $A \subseteq \mathbb{R}^n$ es compacto (esto es, cerrado y acotado) y $f : A \longrightarrow \mathbb{R}$ es continua en A, entonces f alcanza en A el mínimo global y el máximo global.

Punto crítico de un campo escalar.— Sean $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$; se dice que $\vec{a}$ es un punto crítico para f precisamente si $\nabla f(\vec{a})$ no existe o si $\nabla f(\vec{a}) = \vec{0}$.

Condiciones necesarias de primer orden para extremo local.— Si $A \subseteq \mathbb{R}^n$ y $f : A \longrightarrow \mathbb{R}$ presenta un extremo local en $\vec{a} \in A$, entonces $\vec{a}$ es un punto crítico de f. En particular, si f es diferenciable en $\vec{a}$ y presenta un extremo local en $\vec{a}$, entonces $\nabla f(\vec{a}) = \vec{0}$.

Localización de extremos absolutos.— Si $A \subseteq \mathbb{R}^n$ y $f : A \longrightarrow \mathbb{R}$, los posibles puntos en los que f pudiese presentar un extremo absoluto son: bien puntos de la frontera de A, bien puntos interiores de A en los que ∇f no existe, bien puntos interiores de A en los que $\nabla f(\vec{a}) = \vec{0}$.

Punto de silla de un campo escalar.— Punto crítico para un campo escalar donde éste no alcanza un extremo local. En otras palabras, siendo $A \subseteq \mathbb{R}^n$, $f : A \longrightarrow \mathbb{R}$ y $\vec{a} \in A$ un punto no aislado de A, se dice que f alcanza/presenta/tiene un punto de silla en $\vec{a}$ precisamente si para toda bola abierta B centrada en $\vec{a}$ existen puntos $\vec{x} \in B \cap A$ tales que $f(\vec{x}) < f(\vec{a})$ y existen puntos $\vec{y} \in B \cap A$ tales que $f(\vec{y}) > f(\vec{a})$.

Condiciones de segundo orden para extremos locales y puntos de silla.— Sean $A \subseteq \mathbb{R}^n$, A abierto, $f : A \longrightarrow \mathbb{R}$ un campo escalar de clase C^2 en A y $\vec{a} \in A$ un punto crítico de f. Entonces se tienen las siguientes condiciones de segundo orden.

- Condiciones necesarias:
  - Si f presenta un mínimo local en $\vec{a}$, entonces la función hessiana $q_{Hf(\vec{a})}$ es semidefinida positiva.
  - Si f presenta un máximo local en $\vec{a}$, entonces la función hessiana $q_{Hf(\vec{a})}$ es semidefinida negativa.
- Condiciones suficientes:
  - Si la función hessiana $q_{Hf(\vec{a})}$ es definida positiva, entonces f presenta un mínimo local estricto en $\vec{a}$.
  - Si la función hessiana $q_{Hf(\vec{a})}$ es definida negativa, entonces f presenta un máximo local estricto en $\vec{a}$.
  - Si la función hessiana $q_{Hf(\vec{a})}$ es indefinida, entonces f presenta un punto de silla en $\vec{a}$.

Condiciones suficientes de segundo orden para extremos locales y puntos de silla. Caso de dos variables.— Sean $A \subseteq \mathbb{R}^2$, A abierto, $f : A \longrightarrow \mathbb{R}$ un campo escalar de clase C^2 en A y $\vec{a} \in A$ un punto crítico de f. Entonces se satisface:

- Si $\left|Hf(\vec{a})\right| > 0$ y $\dfrac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$, entonces f presenta un mínimo local estricto en $\vec{a}$.
- Si $\left|Hf(\vec{a})\right| > 0$ y $\dfrac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$, entonces f presenta un máximo local estricto en $\vec{a}$.
- Si $\left|Hf(\vec{a})\right| < 0$, entonces f presenta un punto de silla en $\vec{a}$.
- Si $\left|Hf(\vec{a})\right| = 0$, entonces no puede afirmarse nada acerca de lo que f presenta en $\vec{a}$ (habríamos de estudiar el comportamiento de f en un entorno de $\vec{a}$).

Método de los multiplicadores de Lagrange (ML)

ML(n,1). Problema con n variables y una condición

Queremos optimizar el campo escalar $f(\vec{x})$ (*función objetivo*) sujeto a la *condición/restricción* $g(\vec{x}) = 0$. Puede demostrarse que los candidatos a extremos condicionados están entre los puntos críticos de la función auxiliar

$$\Lambda(\vec{x}, \lambda) = f(\vec{x}) - \lambda g(\vec{x}) \quad \text{(lagrangiana)},$$

y que dichos puntos críticos $(\vec{a}, \beta)$ son soluciones de

$$\nabla\Lambda(\vec{x}, \lambda) = \vec{0},$$

esto es, del sistema

$$\begin{cases} \dfrac{\partial\Lambda}{\partial x_1}(\vec{x}, \lambda) &= 0 \\ \dfrac{\partial\Lambda}{\partial x_2}(\vec{x}, \lambda) &= 0 \\ &\vdots \\ \dfrac{\partial\Lambda}{\partial x_n}(\vec{x}, \lambda) &= 0 \\ \dfrac{\partial\Lambda}{\partial \lambda}(\vec{x}, \lambda) &= 0 \end{cases}$$

(Dado el punto crítico $(\vec{a}, \beta)$ de la lagrangiana, $\vec{a}$ es el posible punto donde f alcanza un extremo y β su multiplicador asociado).

Vía 0.ª.—

1. Construimos la lagrangiana $\Lambda(\vec{x}, \lambda)$.
2. Resolviendo $\nabla\Lambda(\vec{x}, \lambda) = \vec{0}$, hallamos los puntos críticos de la lagrangiana, $(\vec{a}, \beta)$. Éstos determinan los puntos $\vec{a}$ posibles extremos de f.
3. Comprobamos que $\nabla g(\vec{a}) \neq \vec{0}$ para cada punto $\vec{a}$ posible extremo de f.
4. Construimos la matriz hessiana de la lagrangiana respecto de las variables x_1, x_2, …, x_n, de orden $n \times n$,

$$H_{\vec{x}}\Lambda(\vec{x}, \lambda) = Hf(\vec{x}) - \lambda Hg(\vec{x}).$$

5. Para cada punto crítico de la lagrangiana, $(\vec{a}, \beta)$:
   I. calculamos el signo de la función hessiana $q_{H_{\vec{a}}\Lambda(\vec{a},\beta))}(\vec{x})$ en aquellos puntos $\vec{x}^* \neq \vec{0}$ tales que $\nabla g(\vec{a}) \cdot \vec{x}^* = 0$;
   II. clasificamos el extremo de f:
      - si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}(\vec{x}^*) < 0$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $g(\vec{x}) = 0$;
      - si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}(\vec{x}^*) > 0$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $g(\vec{x}) = 0$;

- si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}(\vec{x}^*) = 0$, no podemos precisar nada.

Vía 1.ª.—

Los pasos 1, 2, 3 y 4 son los mismos.

5. Para cada punto crítico de la lagrangiana, $(\vec{a}, \beta)$:

   I. calculamos el tipo[5] de la función hessiana $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}$;

   II. clasificamos el extremo de f:

   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}$ es definida negativa, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $g(\vec{x}) = 0$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}$ es definida positiva, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $g(\vec{x}) = 0$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}$ es indefinida, no podemos precisar nada.

Vía 2.ª.—

Los pasos 1, 2, 3 y 4 son los mismos.

5. Para cada punto crítico de la lagrangiana, $(\vec{a}, \beta)$:

   I. construimos la matriz hessiana orlada de la lagrangiana en $(\vec{a}, \beta)$, de orden $(n+1) \times (n+1)$,

$$H\Lambda(\vec{a},\beta) = \left(\begin{array}{c|c} 0 & \nabla g(\vec{a}) \\ \hline \nabla g(\vec{a})^t & H_{\vec{a}}\Lambda(\vec{a},\beta) \end{array}\right)$$

   II. calculamos el signo de cada uno de los menores principales M_k con $k = 3, 4, \ldots, n+1$, donde n es el número de variables;

   III. clasificamos el extremo de f:

   - si todos estos menores son negativos, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $g(\vec{x}) = 0$;
   - si todos estos menores son de signos alternos, siendo $M_3 > 0$ (y, por tanto, $M_4 < 0$, $M_5 > 0$, $\ldots$), entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $g(\vec{x}) = 0$.

ML(n,m). Problema con n variables y m condiciones ($m < n$)

Queremos optimizar el campo escalar $f(\vec{x})$ (*función objetivo*) sujeto a las *condiciones/restricciones* $g_1(\vec{x}) = 0$, $g_2(\vec{x}) = 0$, …, $g_m(\vec{x}) = 0$. Puede demostrarse que los candidatos a extremos condicionados están entre los puntos críticos de la función auxiliar

$$\Lambda(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \lambda_1 g_1(\vec{x}) - \lambda_2 g_2(\vec{x}) - \cdots - \lambda_m g_m(\vec{x}) \quad \text{(lagrangiana).}$$

y que dichos puntos críticos son soluciones de

$$\nabla\Lambda(\vec{x}, \vec{\lambda}) = \vec{0},$$

[5] Si $q_{H_{\vec{a}}\Lambda(\vec{a},\beta)}$ es definida positiva (resp., definida negativa), también será definida positiva (resp., definida negativa) su restricción al subespacio de aquellos puntos $\vec{x}^* \neq \vec{0}$ tales que $\nabla g(\vec{a}) \cdot \vec{x}^* = 0$.

esto es, del sistema

$$\left\{\begin{array}{rcl}\dfrac{\partial\Lambda}{\partial x_1}(\vec{x},\vec{\lambda}) & = & 0\\ \dfrac{\partial\Lambda}{\partial x_2}(\vec{x},\vec{\lambda}) & = & 0\\ & \vdots & \\ \dfrac{\partial\Lambda}{\partial x_n}(\vec{x},\vec{\lambda}) & = & 0\\ \dfrac{\partial\Lambda}{\partial \lambda_1}(\vec{x},\vec{\lambda}) & = & 0\\ \dfrac{\partial\Lambda}{\partial \lambda_2}(\vec{x},\vec{\lambda}) & = & 0\\ & \vdots & \\ \dfrac{\partial\Lambda}{\partial \lambda_m}(\vec{x},\vec{\lambda}) & = & 0\end{array}\right.$$

y que cada punto crítico $(\vec{a},\vec{\beta})$ de la lagrangiana determina un posible punto $\vec{a}$ donde f alcanza un extremo, para lo que es condición suficiente, pero no necesaria, que la matriz jacobiana del campo vectorial de restricciones $\vec{g}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$ (con funciones componentes $g_1,g_2,\ldots,g_m:\mathbb{R}^n\longrightarrow\mathbb{R}$) evaluada en $\vec{a}$, $J\vec{g}(\vec{a})$, tenga rango m (esto es, que en $J\vec{g}(\vec{a})$ exista un menor no nulo de orden m), o lo que es equivalente, que los vectores gradientes $\nabla g_1(\vec{a})$, $\nabla g_2(\vec{a})$, …, $\nabla g_m(\vec{a})$ sean linealmente independientes. A esta condición suficiente se le conoce como *condición de regularidad*.

Vía 0.ª—

1. Construimos la lagrangiana $\Lambda(\vec{x},\vec{\lambda})$.

2. Resolviendo $\nabla\Lambda(\vec{x},\vec{\lambda})=\vec{0}$, hallamos los puntos críticos de la lagrangiana, $(\vec{a},\vec{\beta})$. Éstos determinan los puntos $\vec{a}$ posibles extremos de f.

3. Comprobamos la condición de regularidad para cada punto $\vec{a}$ posible extremo de f.

4. Construimos la matriz hessiana de la lagrangiana respecto de las variables x_1, x_2, …, x_n, de orden $n\times n$,

$$H_{\vec{x}}\Lambda(\vec{x},\vec{\lambda})=Hf(\vec{x})-\sum_{i=1}^{m}\lambda_i Hg_i(\vec{x}).$$

5. Para cada punto crítico de la lagrangiana, $(\vec{a},\vec{\beta})$:

   I. calculamos el signo de la función hessiana $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta}))}(\vec{x})$ en aquellos puntos $\vec{x}^*\neq\vec{0}$ tales que $J\vec{g}(\vec{a})\vec{x}^*=\vec{0}$, esto es, tales que

$$\begin{pmatrix}\dfrac{\partial g_1}{\partial x_1}(\vec{a}) & \dfrac{\partial g_1}{\partial x_2}(\vec{a}) & \ldots & \dfrac{\partial g_1}{\partial x_n}(\vec{a})\\ \dfrac{\partial g_2}{\partial x_1}(\vec{a}) & \dfrac{\partial g_2}{\partial x_2}(\vec{a}) & \ldots & \dfrac{\partial g_2}{\partial x_n}(\vec{a})\\ \vdots & \vdots & \ddots & \vdots\\ \dfrac{\partial g_m}{\partial x_1}(\vec{a}) & \dfrac{\partial g_m}{\partial x_2}(\vec{a}) & \ldots & \dfrac{\partial g_m}{\partial x_n}(\vec{a})\end{pmatrix}\begin{pmatrix}x_1^*\\ x_2^*\\ \vdots\\ x_n^*\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix},$$

   en otras palabras, tales que $\nabla g_1(\vec{a})\cdot\vec{x}^*=0$, $\nabla g_2(\vec{a})\cdot\vec{x}^*=0$, …, $\nabla g_m(\vec{a})\cdot\vec{x}^*=0$;

   II. clasificamos el extremo de f:

   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}(\vec{x}^*)<0$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $\vec{g}(\vec{x})=\vec{0}$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}(\vec{x}^*)>0$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $\vec{g}(\vec{x})=\vec{0}$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}(\vec{x}^*)=0$, no podemos precisar nada.

Vía 1.ª—

Los pasos 1, 2, 3 y 4 son los mismos.

5. Para cada punto crítico de la lagrangiana, $(\vec{a}, \vec{\beta})$:

   I. calculamos el tipo[6] de la función hessiana $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}$;

   II. clasificamos el extremo de f:

   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}$ es definida negativa, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $\vec{g}(\vec{x}) = \vec{0}$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}$ es definida positiva, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $\vec{g}(\vec{x}) = \vec{0}$;
   - si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}$ es indefinida, no podemos precisar nada.

Vía 2.ª—

Los pasos 1, 2, 3 y 4 son los mismos.

5. Para cada punto crítico de la lagrangiana, $(\vec{a}, \vec{\beta})$:

   I. construimos la matriz hessiana orlada de la lagrangiana en $(\vec{a}, \vec{\beta})$, de orden $(m+n) \times (m+n)$,

$$H\Lambda(\vec{a},\vec{\beta}) = \left(\begin{array}{ccc|c} 0 & \dots & 0 & \\ \vdots & & \vdots & J\vec{g}(\vec{a}) \\ 0 & \dots & 0 & \\ \hline & J\vec{g}(\vec{a})^t & & H_{\vec{a}}\Lambda(\vec{a},\vec{\beta}) \end{array}\right)$$

   II. calculamos el signo de cada uno de los menores principales M_k con $k = 2m+1, 2m+2, \dots, m+n$;

   III. clasificamos el extremo de f:

   - si todos estos menores son del mismo signo que $(-1)^m$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un mínimo local estricto condicionado a $\vec{g}(\vec{x}) = \vec{0}$;
   - si todos estos menores son de signos alternos, siendo M_{2m+1} del mismo signo que $(-1)^{m+1}$, entonces $f(\vec{x})$ alcanza en $\vec{a}$ un máximo local estricto condicionado a $\vec{g}(\vec{x}) = \vec{0}$.

Para saber más

UÑA JUÁREZ, Isaías, SAN MARTÍN MORENO, Jesús, y TOMEO PERUCHA, Venancio, *Problemas resueltos de cálculo en varias variables*, Madrid, Madrid (ES-MD), España: Thomson, 2007, ISBN: 978-84-9732-290-4.

[6] Si $q_{H_{\vec{a}}\Lambda(\vec{a},\vec{\beta})}$ es definida positiva (resp., definida negativa), también será definida positiva (resp., definida negativa) su restricción al subespacio de aquellos puntos $\vec{x}^* \neq \vec{0}$ tales que $J\vec{g}(\vec{a})\vec{x}^* = \vec{0}$.